A Novel Earth Mover's Distance Methodology for Image Matching with Gaussian Mixture Models

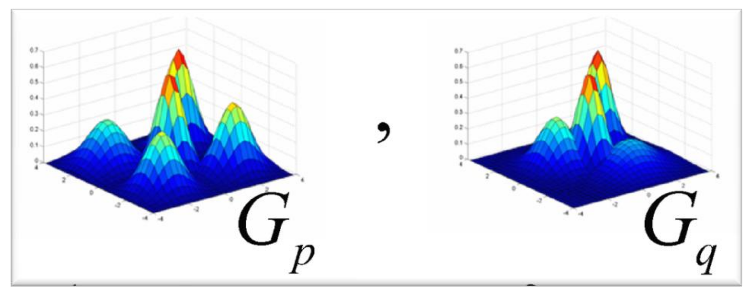
Peihua Li[1], Qilong Wang[2], Lei Zhang[3]

[1]Dalian University of Technology, [2]Heilongjiang University, [3]The Hong Kong Polytechnic University

Introduction

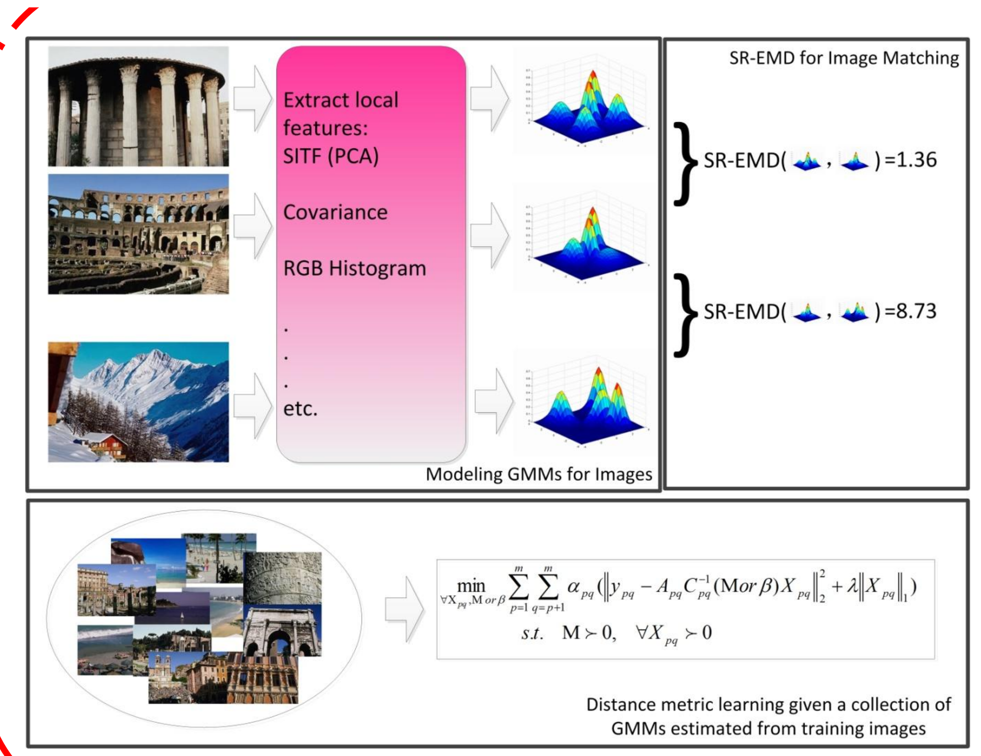
Problem

How to measure dissimilarity/affinity between two Gaussian Mixture Models (GMMs)?



- K-L divergence based ones (Match-KL) [Goldberger et al. ICCV 03]
- Gaussian Quadratic Form Distance (GQFD) [Beecks et al. ICCV 11].
- Earth Mover's Distance (EMD) [Rubner et al. IJCV 00, Logan et al. ICML 01].

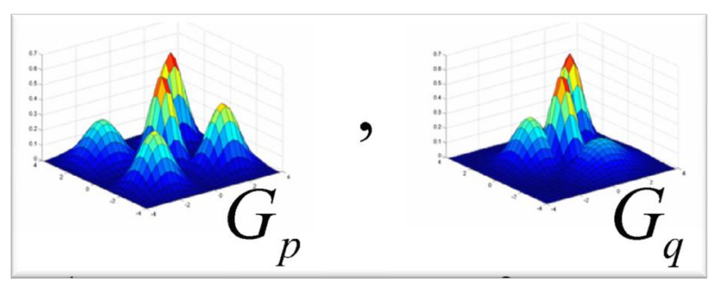
Motivating Applications



Dissimilarity between images is measured by SR-EMD, which is applied to image matching, e.g. image retrieval or texture classification via K Nearest Neighbor (KNN) or Support Vector Machine (SVM)

Proposed EMD Methodology

Given two GMMs:



$G_p = \sum_{i=1}^{n_p} w_i(\boldsymbol{\mu}_i^p, \boldsymbol{\Sigma}_i^p)$ $G_q = \sum_{i=1}^{n_q} w_i(\boldsymbol{\mu}_i^q, \boldsymbol{\Sigma}_i^q)$

Dissimilarity →

Classical EMD is a Transportation problem (TP) ;its cost is $O(n_l(n_p+n_q)^2 n_p n_q))$

$$\mathrm{EMD}(G_p, G_q) = \min_{\mathbf{z}_{pq}} \mathbf{c}_{pq}^T \mathbf{z}_{pq},$$
$$\text{s.t.} \quad \mathbf{A}_{pq}\mathbf{z}_{pq} = \mathbf{y}_{pq}, \ \mathbf{z}_{pq} \succcurlyeq 0$$

$\mathbf{y}_{pq} = [w_1^p, \ldots, w_{n_p}^p, w_1^q, \ldots, w_{n_q}^q], \mathbf{c}_{pq}^T = \{d_{ij}^{pq}\}$

$\mathbf{x}_{pq} = \mathbf{C}_{pq}^T \mathbf{z}_{pq}$ ⇓ $\mathbf{C} = \mathrm{diag}\{d_{ij}^{pq}\}$

SR-EMD

The solution to TP is sparse (#Non-zero entry $\le (n_p+n_q)/(n_p n_q)$), enabling sparse representation-based EMD.

- Tolerable to noise—constraint equations not exactly satisfied
- Low complexity—its cost $O((n_p+n_q)^3 n_p n_q)), n_l \gg (n_p+n_q)$

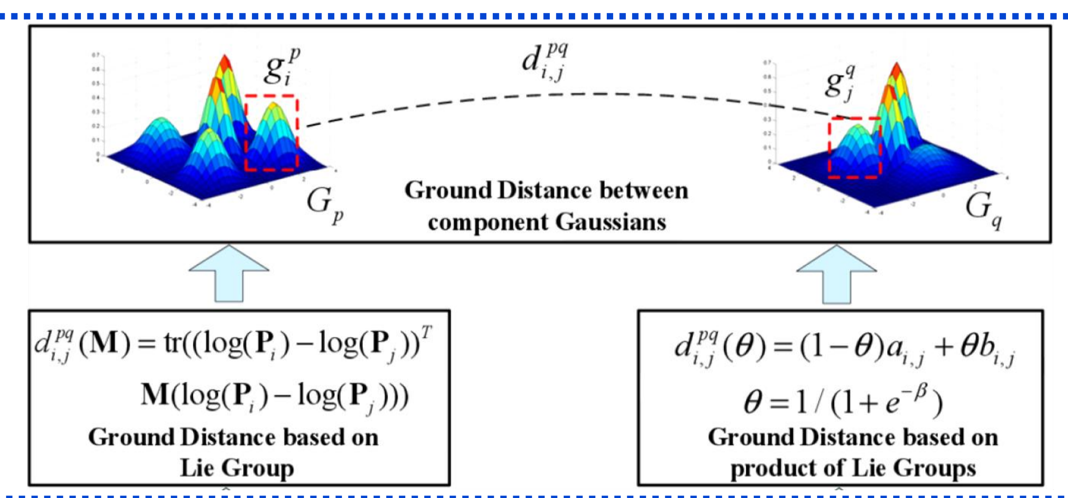
$$\text{SR-EMD}(G_p, G_q) = \min_{\mathbf{x}_{pq}} \frac{1}{2} \| \mathbf{y}_{pq} - \mathbf{A}_{pq}\mathbf{C}^{-1}(\mathbf{M} \, or \, \beta)\mathbf{x}_{pq} \|_2^2 + \lambda \| \mathbf{x}_{pq} \|_1$$
$$\text{s.t.} \quad \mathbf{x}_{pq} \succcurlyeq 0$$

Ground Distances

Ground distances involve distance measures between component Gaussians:

- $d_{ij}^{pq}(\mathbf{M})$ — Embedding Gaussians into the space of Symmetric Positive Definite (SPD) matrices
- $d_{ij}^{pq}(\theta)$ — Product of Lie group

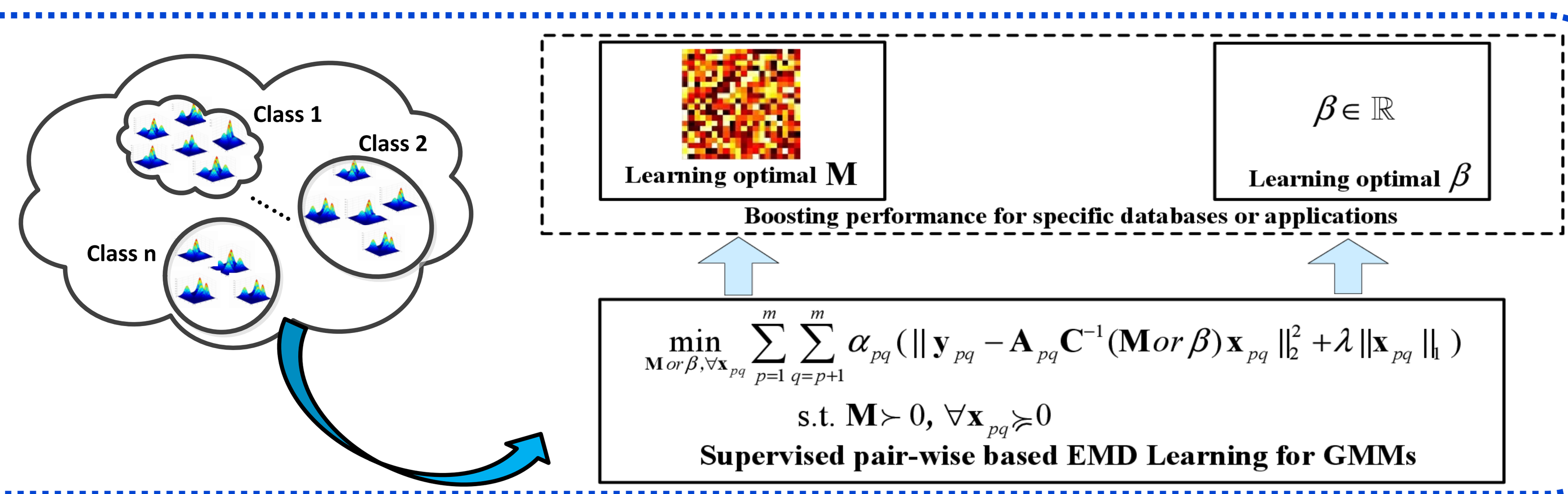
Benefits: Respecting Riemannian structure of the space of Gaussians & facilitating metric learning.



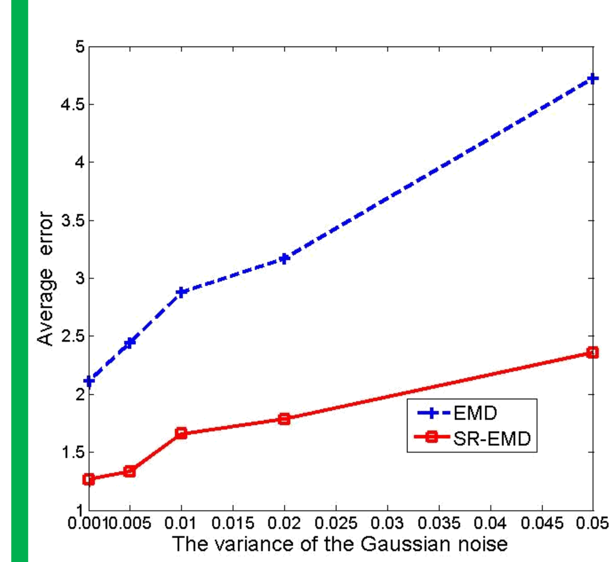
Distance Metric Learning

A simple yet effective pair-wise method for SR-EMD learning given a collection of GMMs. The objective is optimized by alternating two steps:

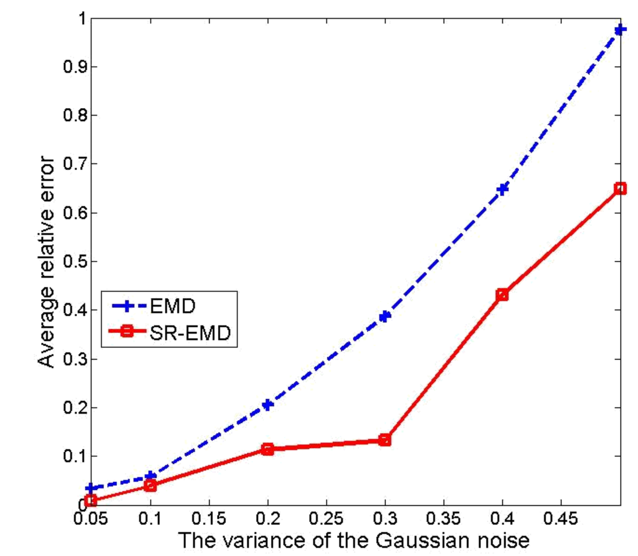
- Solving $\mathbf{x}_{pq}$ by sparse coding method;
- Gradient descent method for $\mathbf{M}$ or θ



SR-EMD vs. EMD



(a) Similar "GMMs"



(b) Dissimilar "GMMs"

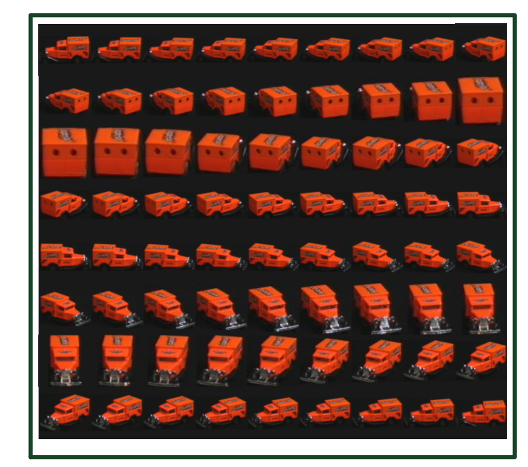
The robustness of SR-EMD and conventional EMD to noise

# of bins	10	20	30	40	50
EMD	6.3e-4	1.0e-2	5.5e-2	2.0e-1	5.6e-1
SR-EMD	5.9e-4	4.7e-3	1.8e-2	4.7e-2	1.2e-1
# of bins	**60**	**70**	**80**	**100**	**120**
EMD	1.28	2.80	5.24	14.7	23.7
SR-EMD	0.23	0.40	0.78	1.79	4.02

Running time comparison between SR-EMD and conventional EMD (unit: s)

Image Retrieval

Corel Wang



Coil100

Method	Match-KL	GQFD	EMD-KL	SR-EMD-θ	SR-EMD-M
Times	6.99	8.70	10.22	5.55	7.63

Running time of different methods (ms)

	Corel WANG			Coil100		
Method	SIFT	Cov	Hist	SIFT	Cov	Hist
Match-KL	40.0	36.7	36.9	28.3	36.1	41.1
GQFD	45.7	46.7	36.8	45.0	47.8	60.0
BoW	46.3	-	-	44.2	-	-
EMD-KL	45.1	38.2	37.1	28.2	39.3	59.7
SR-EMD-M	48.7	48.3	46.9	48.4	61.7	81.8
SR-EMD-θ	49.7	51.6	45.5	52.2	64.2	82.0
SR-EMD-M (with learning)	50.1	51.0	48.7	51.4	64.5	84.9
SR-EMD-θ (with learning)	**52.6**	**53.0**	**48.9**	**55.4**	**69.8**	**85.5**

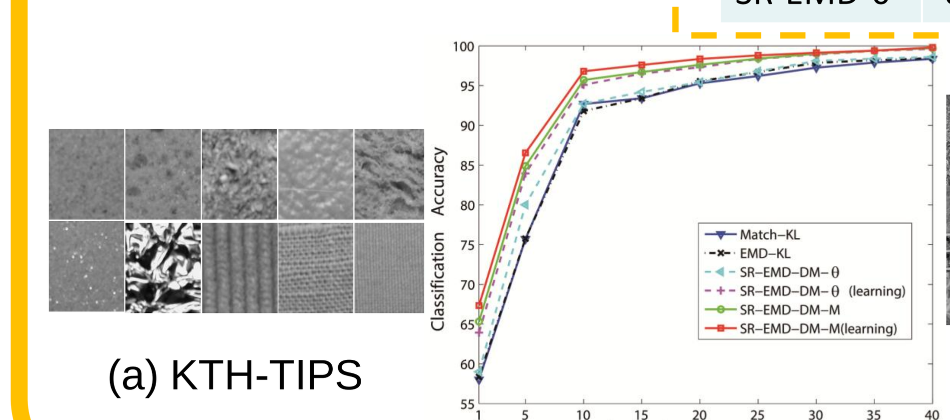
Comparison of MAP values (%)

Texture Classification

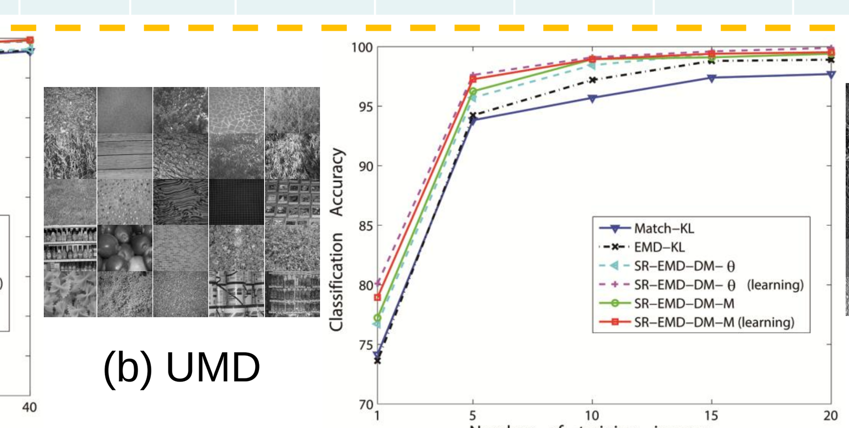
Comparison of texture classification rates (%) with state-of-the-arts

	KTH-TIPS				UMD				CUReT			
Method	1	5	10	40	1	5	10	20	2	10	26	46
Zhang *et al.*	55.1	80.1	90.0	96.1	-	-	-	-	53.6	80.0	91.1	95.3
Hayman *et al.*	50.2	78.3	85.3	94.8	-	-	-	-	60.2	91.0	97.6	98.5
Z-joint	50.5	72.9	80.5	92.1	-	-	-	-	54.4	83.4	93.1	97.4
WMFS	-	-	-	96.5	-	-	-	98.7	-	-	-	-
Liu *et al.*	56.5	80.5	87.8	99.3	73.9	95.0	97	99.3	68.2	91.5	98.3	99.4
CLBP	49.0	76.1	85.5	96.8	73.6	92.4	96.0	98.0	60.2	83.6	92.9	95.9
SR-EMD-M	**67.3**	**86.5**	**96.8**	**99.8**	78.9	97.3	98.4	99.5	71.8	95.2	98.3	99.5
SR-EMD-θ	63.9	84.0	95.1	99.6	**80.1**	**97.6**	**99.1**	**99.9**	**72.1**	**95.4**	**98.7**	**99.5**

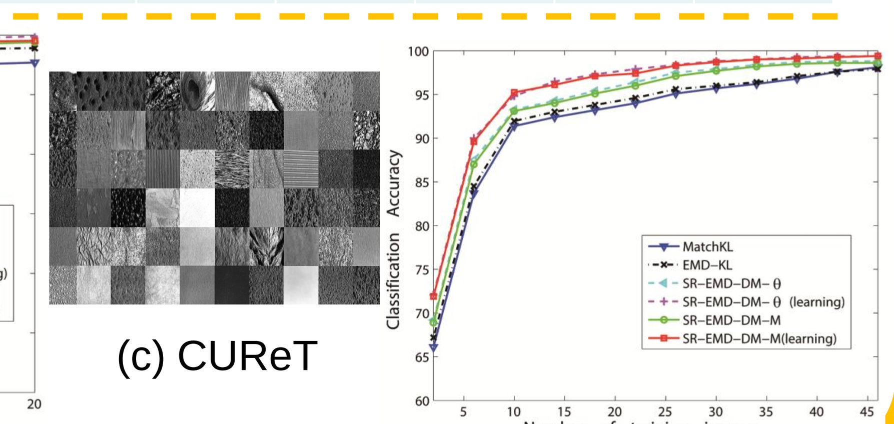
Comparison with various dissimilarity measures



(a) KTH-TIPS



(b) UMD



(c) CUReT